International Journal of Latest Trends in Engineering and Technology
Vol.(16)Issue(1), pp.135-141
DOI: http://dx.doi.org/10.21172/1.171.20
e-ISSN:2278-621X

HAND GESTURE DETECTION FOR LEARNER'S LICENSE USING CNN

Malavika Sreedhar Tankala[1], Manjushri P Rao[1], Sita K Ananda[1], S. Srividhya[2]

Abstract- **In today's world, driving has become an essential part of human lives. In order to legally drive any sort of fuel powered vehicle, there is a need to obtain a permit to do so. There are two steps involved to obtain a driving license. First is to clear the learner's test and then take up the driving test. The proposed work is an application that deals with the process of automating the entire course of the learner's license test.The main working behind this application is using Convolution Neural networks (CNN) which is a part of machine learning. The basic principle behind the working of this application is to provide a computerized test.**

Keywords -Machine learning, Hand gesture, Convolution neural networks.

I. INTRODUCTION

Machine learning in today's world plays a major role in many aspects. It is being extensively made use of in various fields. One of the most important applications of machine learning can be grouped under neural networks. Neural networks are decision making tools that can be used to model complex relationships between inputs and outputs. These neural networks form the basis of the application "Hand gesture detection for learner's license" [1].

To avoid the involvement of a physical authority to be present during the course of the learner's license test, this application has been proposed. Once the candidate enters the test room, the candidate is made to sit in front of a clear background with no disturbances before the test starts. Once the test starts, the candidate is presented with a set of options which state which gesture to present to the system. Once a gesture is selected, the candidate is required to provide the mentioned gesture in the provided frame. It is the responsibility of the candidate to make sure that the gesture is not presented anywhere but the provided frame. Once this is done the candidate can go ahead and take up the test for the rest of the gestures and once all the gestures are covered the candidate can choose to end the test [2].

In this application, the gestures are captured via a web-camera. Once the image is captured, the captured image needs to be classified as whether true or false. True meaning the correct gesture has been provided and false meaning an incorrect gesture has been provided. This image classification is done with the help of preloaded data sets. There are various steps that need to be followed in order to classify the gestures as true or false. Some of the techniques that have been used for this are pre-processing, palm point detection and finger point detection [3].

In the proposed application, the gestures that have been asked to be presented by the candidate are at the moment limited to static gestures only. The four gestures that are to be presented are stop signal, right signal, U-turn signal and slow down signal.

In the following section the papers that were used for reference have been enlisted along with the part of the content that was taken into account for the proposed work. In section 3 a brief description of the implementation has been provided followed by the various tools that have been used for implementation and the results in sections 4 and 5 respectively.

[1] *Department of Information Science and Engineering. BNM Institute of Technology, Bangalore, Karnataka, India*
[2] *Assistant Professor,Department of Information Science and Engineering, BNM Institute of Technology, Bangalore, Karnataka, India*

II. LITERATURE SURVEY

Several new studies and projects have emerged in recent years to apply the concepts of Machine Learning in different fields.

Shivendra Singh and et al, [4] proposed a Real Time Hand Gesture Recognition Using Finger Tips in which the hand region gets extracted from the background by skin color threshold based on the HSI values of the skin. The palm center as well as the finger tips are located right after segmentation of the finger-palm section.Boundary contour is calculated to segment the hand. Tip of the finger is denoted as the peak. Distance transform was used to find the center of the palm using City Block Distance. A rule-based classifier makes use of the number of finger tips detected and the calculated angle to predict the label of hand gesture. To detect the hand, an HSI skin color model is used where this method is susceptible to the intensity of the light. Mohamed Alsheakhali and et al, [5] proposed a Hand Gesture Recognition System which begins by detecting the hand and determining the center, tracking the hands trajectory and finally recognizing the gesture. This technique depends on the following approach which includes different phases. The acquisition phase was where a frame from the webcam is captured followed by segmentation and detection where the image gets segmented into two parts, both of them are used to analyze information such as resultant data, skin pixel values and patterns. A new image gets created containing the location of the center of the hand. Then 10 latest consecutive frames are tracked continuously, in each frame the centers of the moving hands are detected. Then through the user's hands motion, the features are compared with those stored in the database, the maximum likelihood correspondence is chosen. The proposed work is limited since its implemented using RGB values, which requires more processing time. AsanterabiMalima and et al, [6] proposed A Fast Algorithm for Vision-Based Hand Gesture Recognition for Robot Control which includes steps for segmenting the hand region, locating the fingers, and finally classifying the gesture. This method of hand gesture recognition consists of Localizing hand-like regions based on learned skin color statistics producing a Black and White image output. Then performing region-based segmentation of the hand, eliminating small false-alarm regions that were declared as "hand-like" based on their color statistics. Followed by calculating the center of gravity (COG) of the hand region as well as the farthest distance in the hand region from the COG and constructing a circle centered at the COG that intersects all the fingers that are active in the count. Further extracting a 1D binary signal by following the circle, and classifying the hand gesture based on the number of active regions (fingers) in the 1D signal. The proposed work is limited because the hand detection was done based on the Red/Green ratio. Nidhibahen Patel and et al, [7] proposed a Survey on Hand Gesture Recognition Techniques, Methods and Tools under vision-based gesture technique which includes the processes such as taking pictures to make an interface between human and PCs. The acquired pictures are additionally prepared and broken down by utilizing the vision-based procedures. The advantage of this approach was that there was a direct connection and interaction to human and computer devices. The recognition time, strength, speed and effectiveness are challenging issues of this approach. SamataMutha and et al, [8] proposed a Study on Hand Gesture Recognition which includes different techniques and devices that can be used to capture hand gestures. It was also stated that static hand gesture recognition requires training and it has less computational complexity than dynamic hand gesture. One of the most used devices is the depth camera. A depth camera is also called Kinect. Kinect is nothing but RGB-Depth sensor introduced by Microsoft for human computer interaction. In RGB camera only the gesture is recognized whereas in Kinect depth of gesture can be recognized. The main advantage of the Kinect is that the hand gesture is recognized robustly because it can measure the depth. HemlataChavan and et al, [9] HemlataChavan and et al, have proposed A Review on Hand Gesture Detection Using Combine HSI, YCbCr and Morphological Method with Recognition that includes work that investigates dynamic hand gesture recognition using Conditional Random Field. Using the combination of HSI, YCbCr and morphology instead of only HSI and YCbCr color model use in skin colour segmentation. After skin segmentation contour processing is done. Then hand tracking using centroid is done which is followed by feature extraction. HSI colours are determined as far as hue (H) define colours, Saturation(S), and Intensity value (I) define brightness which are the three characteristics that are seen about colour. It signifies the image precisely with institutive qualities. Therefore, usage of HSI values is advantageous. But the methods described above do not produce the same result every time, errors in edges and boundaries for hand detection, noise, problem over segmentation, memory requirement, background with red colour, complex background, cluttered background, recognition rate and other accuracy problems.

III. IMPLEMENTATION

The architecture of the proposed system in Figure.1.depicts the overall working of the proposed hand gesture detection system.

A. Image capture

Images are captured through a normal camera. The images are taken under ideal condition i.e. controlled illumination. Therefore, it is easy and effective to detect the hand region. The HSI model is used to measure the skin color. The values 20, 255 and 255 are taken as upper HSI values of skin color and 0,50 and 90 are taken as the lower HSI values of the skin color. A filter is used to filter out other values that do not fall in this range [10].

B. Pre-processing

The detected region contains the arm region along with other noises due to color similarity. The Hand Gesture detection for Learner's License vicinity which represents hand has all of the pixel values set as 0 and everything else is considered with pixel intensity set 1. To reduce the background noise, we cast off all significant smudges that are connected to components from the images those have area fewer than P pixels. Boundary contour is calculated to segment the hand. It is achieved by scanning the hand from left to right and then scanning that scanned boundary from top to bottom. The segmentation of palm-finger region from arm region can be done by finding area and length of the contour.

C. Palm point detection

The palm point is known as the center of the palm. Distance transform can be used to find the center of the palm. They can also be called as distance map; distance transform can be done by representing the image based on distance pixels.

In distance transform image, each pixel will calculate the distance of it and the nearest boundary pixel. City Block distance is the method that is used to measure the distance between the pixels and the nearest boundary pixels. In the distance transform image of the binary form, the pixel with the largest value is chosen as the palm point.

D. Finger-tip detection

In this step, tip of the finger is denoted as peak. Based on the number of peaks the total numbers of finger raised can be found. The fact that fingers are of various height is applied. To locate the finger tips, the hand region is split into two vertical regions. The image is scanned for each half. First half of is scanned from left to right, beginning from top and second half from right to left beginning from top. Whenever a pixel of '0' intensity is detected (hand region pixel value set to zero in pre-processing), it is marked as tip. The vertical half wherein the tip is detected is reduced to new size (wherein the tip was detected); in addition, the top-most point is also reset. Again, the image is scanned for new halves and this technique continues. Some false peak can be detected during fingertip detection. These peaks can bring about false classification. To eliminate these undesirable peaks a threshold is computed. Here 0.07 times the most peak value is used as threshold. Values more than this threshold are considered as peak.

E. Gesture classification

Classification of different hand gesture can be done by identifying the number of finger tips detected and the angle between each fingertip and the palm point. For example, if three fingers Hand gesture detection for learner's license are detected, and the angle formed by each finger with the palm point is a1, a2, a3 (matching the training database) only then it will be labelled as three.

F. Result

There are a few constrains and conditions assumed in the proposed work for better results. First is the normal environment- The proposed method works well with static or homogeneous background. They are based on hue, saturation, and intensity values to detect the hand region. The hand detection rate gets effected if there is variation in light intensity. Hence, all training and testing phase was performed in controlled environment and fixed light intensity was used [11].

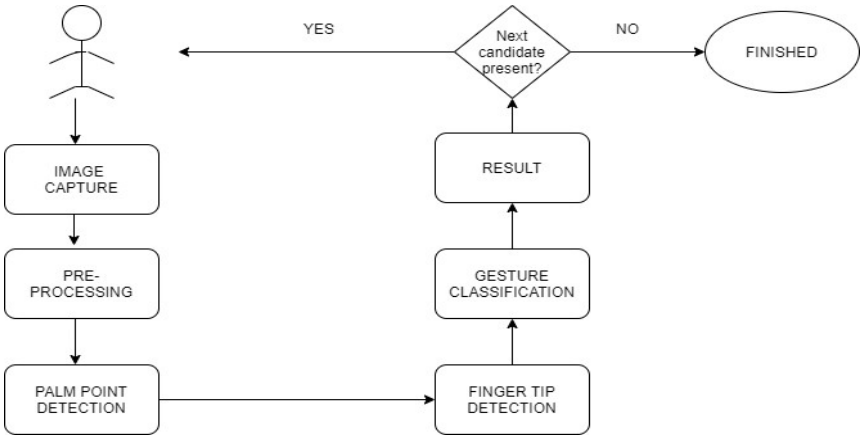

Figure. 1. Implementation Diagram

IV. TECHNOLOGIES USED

A. Convolution Neural Network

CNN is a class of deep, feed – forward artificial neural networks (where connection between nodes do not form a cycle) and use a variation of multilayer perceptron's designed to require minimal pre-processing [12]. CNN Architecture shown in the following Figure.2. uses convolutional land pooling layers. Pooling is a form of non-linear down-sampling and are applied after convolutional layers to remove unnecessary features and decrease the number of parameters during the training in a quest to prevent overfitting [13].

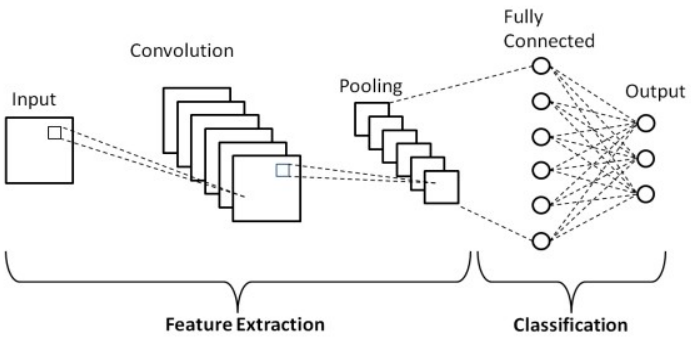

Figure.2. CNN Architecture

Specifically, use of max pooling (MaxPool2D) is done [14]. After the convolutional and max pooling layers, use of the flattening and fully connected (FC) layers, which flatten the multidimensional array into a two-dimensional one [15]. The output layer is based on the SoftMax function, which calculates the probabilities of each class [16] [17].

B. TensorFlow

TensorFlow is one of the open source libraries which helps one to develop and train ML models. They use dataflow graph to build models which allows the developer to visualize the construction of the neural network with Tensor board. TensorFlow can be used for classification, perception, Understanding, Discovering, Prediction and Creation. It permits the developers to create large-scale neural networks with many layers. Tensorflow library includes specific API to build at scale deep studying architecture like CNN or RNN [18].

V. RESULT

The front end contains four gesture buttons and an evaluate button which is shown in Figure.3. The user has to click on each button and show the respective gesture that is mentioned on the button

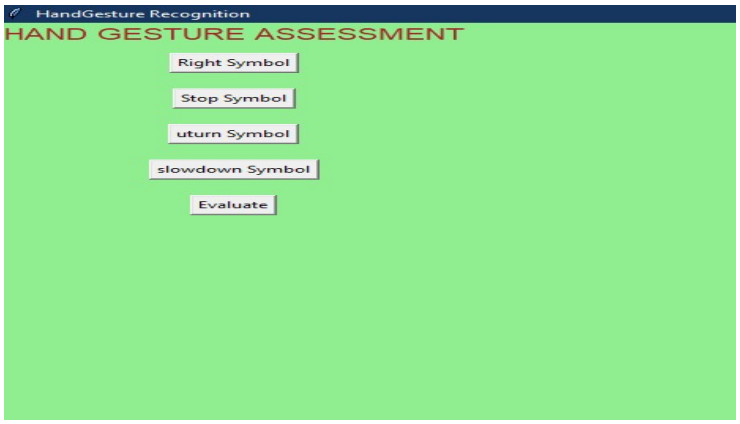

Figure. 3. Front End

Once the user clicks on the right gesture button a camera pops up where the user has to show a right gesture within the frame displaying on the screen which is shown in the Figure.4.Similarly the user have to show stop, uturn and slowdown gesture by clicking the respective hand gesture button which is shown in the Figure. 5,6,7 respectively

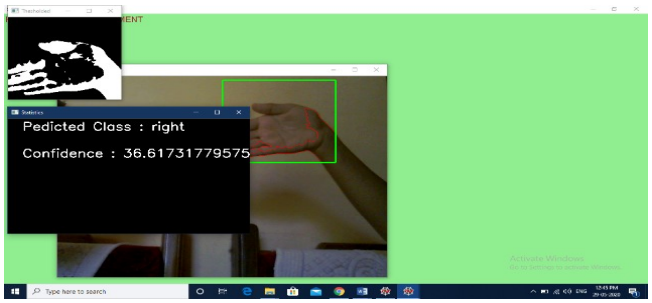

Figure. 4. Right Gesture

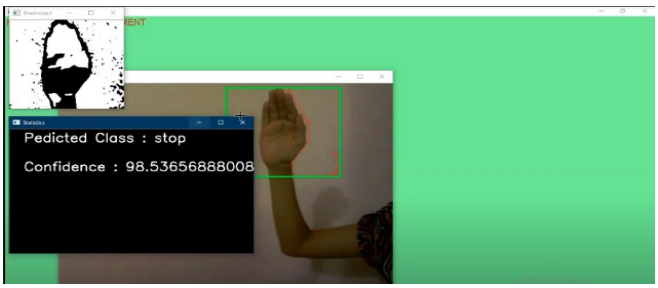

Figure. 5. Stop Gesture

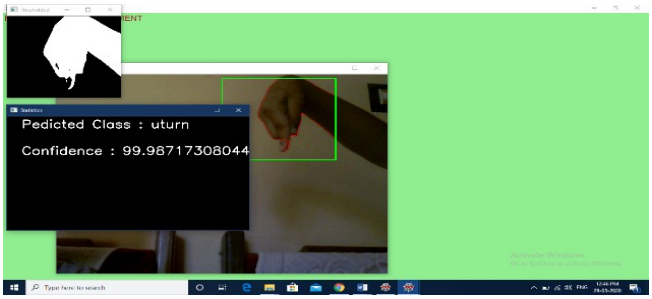

Figure. 6. U-Turn Gesture

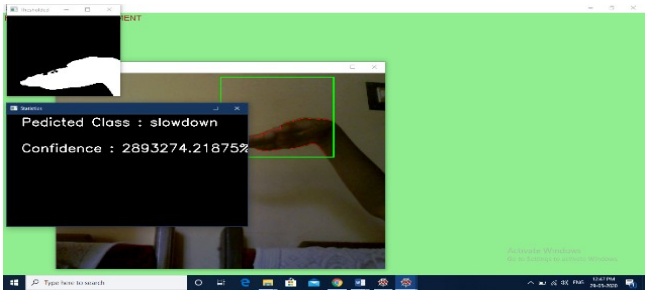

Figure. 7. Slowdown Gesture

When user clicks on the evaluate button the result will get generated showing the predicted output as well as the actual output which is shown in the Figure.8.

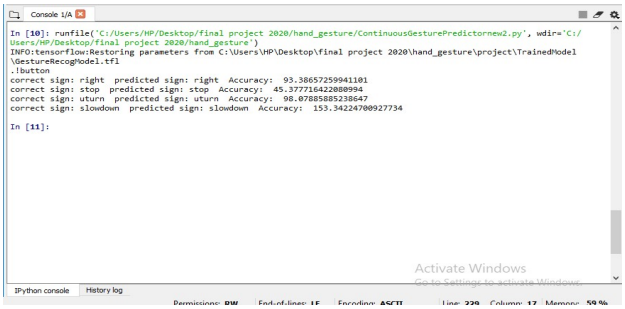

Figure. 8. Accuracy Rate

VI. CONCLUSION

The candidate to take up the leaners license test has to go through a series of tests which include an MCQ as well as a hand gesture detection test. Currently there is a need for an authority to be present at the center where the test is conducted.

The proposed work aims to eradicate the need for an external authority to be physically present at the time of learner's license test.

This proposed application captures a gesture presented by a candidate in a controlled test environment and tries to classify if the projected gesture is the same as what is expected from the candidate.

The main intention of this proposed application is to help automate the entire process of the learner's license test.

REFERENCES

[1] https://pdfs.semanticscholar.org/7785/7eef494b30663e0c17270486f35851246e6d.pdf
[2] http://m-hikari.com/ces/ces2017/ces5-8-2017/p/daiCES5-8-2017.pdf
[3] https://www.ijcsmc.com/docs/papers/January2015/V4I1201506.pdf
[4] Shivendra Singh: "Real Time Hand Gesture Recognition Using Finger Tips", International Research Journal of Engineering and Technology (IRJET) Volume: 05 Issue: 06, pp. 1420-1424, June 2018.
[5] Mohamed Alsheakhali, Ahmed Skaik, Mohammed Aldahdouh, Mahmoud Alhelou: "Hand Gesture Recognition System", Computer Engineering Department, The Islamic University of Gaza Strip, Palestine, Volume: 01, pp. 01-02,2011.

[6] AsanterabiMalima, ErolÖzgür, and MüjdatÇetin: "A Fast Algorithm for Vision-Based Hand Gesture Recognition for Robot Control", Faculty of Engineering and Natural Sciences, Sabancı University, Tuzla, İstanbul, Turkey, Volume: 13, pp. 389-394, 2006.
[7] Nidhibahen Patel, Dr. Selena (Jing) He: "A Survey on Hand Gesture Recognition Techniques, Methods and Tools, International Journal of Research in Advent Technology", International Journal of Research in Advent Technology, Vol.6, No.6, June 2018.
[8] SamataMutha, Dr. K.S. Kinage: "Study on Hand Gesture Recognition", IJCSMC, Vol. 4, Issue. 1, pp. 51 – 57, January 2015.
[9] HemlataChavan, Mr. Prateek Gupta: "A Review on Hand Gesture Detection Using Combine HSI, YCbCr and Morphological Method with Recognition", International Research Journal of Engineering and Technology (IRJET), Volume: 03 Issue: 05, pp. 1039-1042, May2016.
[10] Zhihua Hu, Xiaoming Zhu: "Gesture detection from RCB hand image using modified convolution neural network", School of Computer Science, Huanggang Normal University, IEEE, 2019.
[11] Xing Yingxin, "A Robust Hand Gesture Recognition Method Via Convolutional Neural Network", Beijing Key Laboratory of Multimedia and Intelligent Software Technology College of Metropolitan Transportation Beijing University of Technology Beijing, IEEE, 2016.
[12] Jaya Prakash Sahoo, Samit Ari: "Hand Gesture Recognition using PCA based Deep CNN Reduced Features and SVM classifier", Department of Electronics and Communication Engineering, National Institute of Technology, Rourkela Odisha, IEEE, 2019.
[13] Hung- Yuan Chung: "An Efficient Hand Gesture Recognition System Based on Deep CNN", Department of Electrical Engineering, National Central University Taoyuan, Taiwan, IEEE, 2019.
[14] KolliparaSaiVarun, Puneeth, T. Prem Jacob, "Hand Gesture Recognition and Implementation for Disables using CNN's", International Conference on Communication and Signal Processing, IEEE, April 4, 2019.
[15] Richard Adiguna, YustinusEkoSoelistio,"CNN Based Poster-Free Hand Detection", Information System Dept, Universitas Multimedia Nusantara Indonesia, IEEE, 2018.
[16] ShilpaChaman, Dylan D Souza, Benz D mello, Karan Bhavsar, Jolton D Souza, "Real-time Hand Gesture Communication System in Hindi for Speech and Hearing Impaired", Student, EXTC Department, St. Francis Institute of Technology Mumbai, IEEE,2018.
[17] Saransh Sharma, Samyak Jain, "A Static Hand Gesture and Face Recognition System for Blind People", Khushboo, Dr. Akhilesh Das Gupta Institute of Technology and Management, New Delhi, IEEE, 2019.
[18] Faith Ertam, Galip Aydin, "Data Classification with Deep Learning using Tensorflow", Informatics Department, Firat University, IEEE, 2017.